

NS

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 16:30:34 ; Search time 25 Seconds
(without alignments)
1246.770 Million cell updates/sec

Title: US-09-635-501-2
Perfect score: 4291
Sequence: 1 MSSSSWLLSLVAVTAAQST.....ISKGENNPGFQNTDDVOTSF 805

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1344	31.3	732	1 ACET_HUMAN	P22966 homo sapien
2	1337	31.2	1306	1 ACET_HUMAN	P12821 homo sapien
3	1334	31.1	732	1 ACET_MOUSE	P22967 mus musculus
4	1334	31.1	1312	1 ACET_MOUSE	P09470 mus musculus
5	1312	30.6	1193	1 ACET_CHICK	Q10751 gallus gall
6	1310	30.5	1313	1 ACET_RAT	P47820 rattus norv
7	1283.5	29.9	1310	1 ACET_RABIT	P12822 oryctolagus
8	1283	29.9	737	1 ACET_RABIT	P22968 oryctolagus
9	1090	25.4	615	1 ACET_DROME	Q10714 drosophila
10	1058	24.7	611	1 ACET_HAETE	Q10715 haematobia
11	135	3.1	501	1 YPWA_BACSU	P50848 bacillus su
12	125	2.9	627	1 GLGB_BACSU	P39118 bacillus su
13	123	2.9	986	1 EP1B_STAEP	P30195 staphylococ
14	121	2.8	1034	1 BGAL_BACME	O52847 bacillus me
15	119.5	2.8	3655	1 YAMB_SCHPO	Q10064 schizosacch
16	119.5	2.8	4540	1 DYHC_PARTE	Q27171 paramecium
17	118	2.7	3911	1 AKAG_HUMAN	Q99996 h a-kinase
18	115.5	2.7	3433	1 UTRO_HUMAN	P46939 homo sapien
19	115	2.7	950	1 Y511_RICPR	Q92d36 rickettsia
20	114.5	2.7	901	1 P1P_LACIA	P49022 lactococcus
21	114.5	2.7	1225	1 Y309_MYCGE	P47551 mycoplasma
22	113.5	2.6	6669	1 NEBU_HUMAN	Q20929 homo sapien
23	112.5	2.6	773	1 CDH_PHACH	Q01738 phanerochae
24	109.5	2.6	1398	1 DNA2_SCHPO	Q99u2 schizosacch
25	109	2.5	906	1 Y010_CLOAB	Q97n28 clostridium
26	108.5	2.5	1084	1 XP01_YEAST	P30822 saccharomyc
27	108.5	2.5	2167	1 YCS2_YEAST	P25356 saccharomyc
28	107.5	2.5	1689	1 RPA1_SCHPO	P15398 schizosacch
29	107.5	2.5	3685	1 DMD_HUMAN	P11532 homo sapien
30	107	2.5	736	1 VP4_ROTTC	P26193 porcine rot
31	106.5	2.5	1312	1 RA50_YEAST	P12753 saccharomyc
32	106	2.5	1279	1 BCHH_CHLVI	O50314 chlorobium
33	105.5	2.5	645	1 REP_BUCAL	P57634 buchnera ap

RESULT 1	ACET_HUMAN	STANDARD;	PRT;	732 AA.
ID	ACET_HUMAN			
AC	P22966;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Angiotensin-converting enzyme, testis-specific isoform precursor (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (Kininase II).			
DE	DCPI OR DCP OR ACE.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OC	NCBI_TaxID=9606;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=90046671; PubMed=2554286;			
RX	Ehlers M.R.W., Fox E.A., Strydom D.J., Riordan J.F.;			
RA	"Molecular cloning of human testicular angiotensin-converting enzyme: the testis isozyme is identical to the C-terminal half of endothelial angiotensin-converting enzyme."			
RT	Proc. Natl. Acad. Sci. U.S.A. 86:7741-7745(1989).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=89338720; PubMed=2547653;			
RX	Lattion A.L., Soubrier F., Allegrini J., Hubert C., Corvol P.,			
RA	Alhenc-Gelas F.;			
RT	"The testicular transcript of the angiotensin I-converting enzyme encodes for the ancestral, non-duplicated form of the enzyme."			
RL	FEBS Lett. 252:99-104(1989).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANTS P-32; G-49 AND S-712.			
RX	MEDLINE=99251580; PubMed=10319862;			
RA	Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;			
RT	"Sequence variation in the human angiotensin converting enzyme."			
RL	Nat. Genet. 22:59-62(1999).			
RN	[4]			
RP	ZINC-BINDING.			
RX	MEDLINE=91308093; PubMed=1649623;			
RA	Ehlers M.R., Riordan J.F.;			
RT	"Angiotensin-converting enzyme: zinc- and inhibitor-binding stoichiometries of the somatic and testis isozymes."			
RL	Biochemistry 30:7118-7126(1991).			
CC	-1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.			
CC	-1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide, oligopeptide-I-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither Asp nor Glu. Converts angiotensin I to angiotensin II.			
CC	-1- COFACTOR: BINDS 1 ZINC ION.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME IS TRANSCRIPTED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY FROM AN ALTERNATIVE START SITE.			
CC	-1- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.			
CC	-1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL			

P24740 human immun
P38903 saccharomyc
P30538 bacillus st
Q14511 homo sapien
P25805 plasmodium
Q06366 clostridium
Q08253 homo sapien
P32775 saccharomyc
Q00174 drosophila
Q92h03 streptococ
P05897 simian immu

ALIGNMENTS

REGULATION BY ANDROGENS.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M26657; AAA60611.1; -;
DR EMBL; X16295; CRA34362.1; -;
DR EMBL; AF118569; AAD28561.1; -;
DR PIR; S05238; S05238;
DR PIR; A33979; A33979;
DR MEMOPS; M02.004; -;
DR MIM; 106180; -;
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR00130; Zn_MPeptidse.
DR Pfam; PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPTIDASEA.
DR ProDom; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROPEASE; 1.
KW Hydrolase; Metalloprotease; Carboxypeptidase; zinc; Dipeptidase;
KW Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing;
KW Polymorphism.
FT SIGNAL 1 31 ANGIOTENSIN-CONVERTING ENZYME,
FT CHAIN 32 732 TESTIS-SPECIFIC ISOFORM.
FT DOMAIN 32 684 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 685 701 POTENTIAL.
FT DOMAIN 702 732 CYTOPLASMIC (POTENTIAL).
FT METAL 414 414 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 415 415 BY SIMILARITY.
FT METAL 418 418 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 442 442 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 457 457 POTENTIAL.
FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 32 32 S -> P.
FT VARIANT 49 49 /FTIG=VAR_011710.
FT VARIANT 712 712 /FTIG=VAR_011711.
FT VARIANT 712 712 R -> S.
FT SEQUENCE 732 AA; 83330 MW; 80B0D19CFA642313 CRC64;
Query Match 31.3%; Score 1344; DB 1; Length 732;
Best Local Similarity 41.8%; Pred. No. 7.4e-86;
Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;
YQ 15 TAAQS-----TIERQAKTFDKFHEADELFYQSSLASWYNTNITEE-----NVQNM 62
DB 61 TSAQSNLYTDEAEASKFEVEYDRTSOVVMYAEANWYNITTTETSKILLQKNMQIA 120
YQ 63 NNAGDKWAFKQSTLAQWYPIQETONITVKLQALQOQSSVLSDEKSKRLNTILT 122
DB 121 NHT-----LKYQARKFDVNLQNTTIKRIKKVDLQERAAQAQLEEYNNKILLD 172
YQ 123 MSTYISGKVCNPDNPQECILLPEGLNEIMANSLDYNERLWAEWSRVSQOLRPLYEE 182
DB 173 METTYSVATVCHPNG--SCLOLEFDLTNNATSRKYEDLLWAEGRDRAGRAILOFYFK 230
YQ 183 YVYLKNEAMAHYEDYGDYWRGDYEVNGVDGYSGROLIEVHTFEIKRPLYEHLHA 242
DB 231 YVELINQAARLNGYVDAGDSWRSMYETPSLE-----QDLRLFQELQPLYLNLHA 280
YQ 243 YVRACLNNAY-PSYISPIGCLPAHLGLDMWGRWNTNLSLYVPPFGKPNIDVDMVDQA 301

DB 281 YVRRALHRRHYGQHINLEGPAPAHLLGNMWAQTSNIYDLVWPPSPASMDITTEAMLKQG 340
YQ 302 WDQRIFKEAEKFSVGLPNMTQGEWNSMLTDPGNQKAVCHPTAWDLGKG-DFRILM 360
DB 341 WTPRMFKADDFTSGLGLLPVPEFWNKSLEKPTDGRVNVCHASAWDFYNGKDFRIKQ 400
YQ 361 CTKVTMDFTTAHEMHGHIQYDMAYAAQPLFLNLANGNEGHEAVGEIMLSAATPKHLS 420
DB 401 CTTVNLEDLVAAHEMHGHIQYDMAYAAQPLFLNLANGNEGHEAVGEIMLSAATPKHLS 460
YQ 421 IGLSPDFQEDNETEINFLIKQALITVGTLPFTYMLEKWEWYFKGEIPKDOHKKWEM 480
DB 461 INLLSSEGS-D-BEIDINFLMKMALDKIAFPFVYLDVQWRVDFGSGITENTNQEWNSL 519
YQ 481 KREIVGVPEVPVDETYCDPASLPHVSNDSYFRTYRTLYQFQFQALCOAAHESGPLH 540
DB 520 RLKYQGLCPVPRTOGDFDPCAFHIPPSSVYIRYFVSFIQFQFHEALCOAAHGTGPLH 579
YQ 541 KCDISNSTEAGQKLFNMLRLGKSEPTWTLAENVGAKNNVRPLLNFPELFTWLKDQNK 600
DB 580 KCDIYQSKAQRLATANKLGFSRPWEAMQLITGPNMSASAMLSYFKPLDMLRTENE 639
YQ 601 --NSFVGH-STDSWSPYADQS 617
DB 640 LHGEKLGMPQYNWTPNSARS 659
RESULT 2
ACE_HUMAN STANDARD; PRT; 1306 AA.
ID AC PI2821;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
DE (ACE) (Dipeptidyl carboxypeptidase I) (Kininase II) (CD143 antigen).
GN DCP1 OR DCP OR ACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89071703; PubMed=2849100;
RA Soubrier F., Alhenc-Gelas F., Hubert C., Allegrini J., John M.,
Tregear G., Corbol P.;
RT "Two putative active centers in human angiotensin I-converting enzyme
revealed by molecular cloning."
RL Proc. Natl. Acad. Sci. U.S.A. 85:9386-9390(1988).
RN [2]
RX SEQUENCE FROM N.A., AND VARIANTS S-261; W-561 AND S-1286.
RA MEDLINE=9251580; PubMed=10319862;
RA Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;
RT "Sequence variation in the human angiotensin converting enzyme."
RL Nat. Genet. 22:59-62(1999).
RN [3]
RP PARTIAL SEQUENCE OF 30-46.
RC TISSUE=Lung;
RX MEDLINE=90110025; PubMed=2558109;
RA Takeuchi K., Shimizu T., Ohishi N., Seyama Y., Takaku F.,
Yotsumoto H.;
RT "Purification of human lung angiotensin-converting enzyme by high-
performance liquid chromatography: properties and N-terminal amino
acid sequence."
RL J. Biochem. 106:442-445(1989).
RN [4]
RP ZINC-BINDING
RX MEDLINE=91308093; PubMed=1649623;
RA Ehlers M.R., Riordan J.F.;
RT "Angiotensin-converting enzyme: zinc- and inhibitor-binding
stoichiometries of the somatic and testis isozymes."
RL Biochemistry 30:7118-7126(1991).

CC -|- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -|- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-|-xaa-Abb, when xaa is not pro, and xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -|- COFACTOR: BINDS 2 ZINC IONS.
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -|- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME
 CC IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY
 CC FROM AN ALTERNATIVE START SITE.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC -|- DATABASE: NAME-PROW; NOTE-CD guide CD143 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd143.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J04144; AA51684.1; -.
 CC EMBL: AF118569; AAD28560.1; -.
 CC PIR: A31759; A31759.
 CC PIR: PQ0004; PQ0004.
 CC MIM: 106180; -.
 CC InterPro: IPR001548; Peptidase_M2.
 CC InterPro: IPR000130; Zn_MTPeptide.
 CC Pfam: PF01401; Peptidase_M2; 2.
 CC PRINTS: PR00791; PEPTIDASEA.
 CC ProDom: PD004184; Peptidase_M2; 2.
 CC PROSITE: PS00142; ZINC_PROTEASE; 2.
 CC DR Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing;
 KW Polymorphism.
 FT SIGNAL 1 29
 FT CHAIN 30 1306
 FT
 FT DOMAIN 30 1259 ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
 FT TRANSMEM 1260 1276 ISOFORM.
 FT DOMAIN 1277 1306 POTENTIAL.
 FT REPEAT 227 583 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 825 1181
 FT METAL 390 390 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 391 391 1 (BY SIMILARITY).
 FT METAL 394 394 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 988 988 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 989 989 2 (BY SIMILARITY).
 FT METAL 992 992 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 509 509 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 677 677 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 695 695 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 714 714 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 760 760 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 942 942 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1191 1191 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 261 261 A -> S.
 FT
 FT VARIANT 561 561 /FTid=VAR_011707.
 FT
 FT VARIANT 1286 1286 R -> W.
 FT
 FT VARIANT 1286 1286 R -> S.

FT CONFLICT 35 35 /FTid=VAR_011709.
 FT CONFLICT 42 42 Q -> E (IN REF. 2).
 FT CONFLICT 42 42 D -> R (IN REF. 2).
 SQ SEQUENCE 1306 AA; 149714 MW; 1B33BCA7301A26AA CRC64;
 Query Match 31.2%; Score 1337; DB 1; Length 1306;
 Best Local Similarity 41.7%; Pred. No. 5.2e-85;
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;
 QY 20 TIEQAKTFLDKFNHEADLFYQSSLASNYNTNITEE-----NYQNMMNADGKWSA 71
 DB TDEAEASKFVEEYDRTSQVWVNEAEANWNTNITTETSKILLQKNQJANHIT- 697
 QY 72 FLKEGSTLAQWYPLQEIQLNVLKQLQALQONGSVLSEDKSKRLNTLNTMTSIYSNGK 131
 DB --LKYGTQARKFDVNLQNTTIKRIKKVQDLERAALPAQALEEYNNKILLDMETYSVAT 755
 QY 132 VCNPNPQECILLEPLGMEINMANSIDYNERLWAMESRSEVGKOLRPLYEYVVLKNEMA 191
 DB VCHPNG--SCLEPDLTNVATSKYEDLLWANEGRDKAGRAILQYPKYVELINQAA 813
 QY 192 RANHYEDYGDYWRGDIYVNGVDYDSRGQIEDVDEHTFEIKPLYEHLHAYVRAKLMA 251
 DB RLNGYVDAGDSWRSWYETPSLE-----QDLERLFQELQPLYLNLHAYVRRALHRH 863
 QY 252 Y-PSVISPIGCLPAHLGLGMGRFWTNLYSLTVPGOKPNTDVTDMYDQANDAQRIKE 310
 DB YGAQHINLEGGPIPAHLGLGNMWAQTWSNIYDLVVPFSPASMDTTEAMLKQGTFRMFKE 923
 QY 311 AEKFFVSVGLPNTQGFWENSKLTDPGNVOKAVCAHPTAWDLGK--DFRILMCTKVTWDDF 369
 DB ADDFTSLGLLPVPEPFWPKSKLEKPTDGRVVCVCHASAWDFYNGKDFRIKQCTTVNLEDL 983
 QY 370 LTAHHENGHIQDYMAAAQPFLLRNGANEHGEVHAEVGEINLSAATPHLKGISGLSPDFQ 429
 DB VVAHHENGHIQFYMQYKDLPAVALREGANPGFHEAIGDVLALSVSPTPKHLHLNLLSSEGG 1043
 QY 430 EDNETEINFLKQALTVITGLPTFMLEKRWVMVFKGIPKDONMKWENKREIVGVVE 489
 DB SD-EHDINFLMKALDKIAFPFSYLVDMWRVDFDSITKENTNQENWWSLQKGLCP 1102
 QY 490 PVPHEITYCDPASLFHVSNDYSPIRYTYRTLYQFQFALCQAAKHEGHLKCDISNSTE 549
 DB PVPRTQGDGDPGAKFHPSPSVPIRYFVSFIQFQFHEALCQAAAGHTGPKHKDIYQSKE 1162
 QY 550 AGQKLFNMLRGKSEPTWLALENVVGAKNNVRPLLHYFFELFWLKDQNK--NSFVGVW- 606
 DB AGQRLATAMKLGFRPWPPEAMQLITGPNNMSASAMLSYFKRPLLDLWRTENELHGERLWGP 1222
 QY 607 STDWSPVADQS 617
 DB QYNWTPNSARS 1233
 RESULT 3
 ACET_MOUSE
 ID ACET_MOUSE STANDARD; PRT; 732 AA.
 AC P22967;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Angiotensin-converting enzyme, testis-specific isoform precursor
 DE (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase 1) (Kininase II).
 GN DCP1 OR ACE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90318396; PubMed=2164636;
 RA Howard T.E., Shai S.-Y., Langford K.G., Martin B.M., Bernstein K.E.;
 RT "Transcription of testicular angiotensin-converting enzyme (ACE) is

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; J04946; AAA37147.1; -
 CC EMBL; J04947; AAA37148.1; -
 CC EMBL; J03940; AAA37146.1; -
 CC PIR; A29220; A29220.
 CC PIR; A34171; A34171.
 CC MGD; MGI:87874; Ace.
 DR InterPro; IPR001548; Peptidase_M2.
 DR Pfam; PF01401; Peptidase_M2; 2.
 DR PRINTS; PR00791; PEPDIPASEA.
 DR PROSITE; PD004184; Peptidase_M2; 2.
 DR PROSITE; PS00142; ZINC_PROTEASE; 2.
 KW Hydrolase; Metalloprotease; Carboxypeptidase; zinc; dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
 FT SIGNAL 1 34
 FT CHAIN 35 1312
 FT
 FT DOMAIN 35 1264
 FT TRANSMEM 1265 1281
 FT DOMAIN 1282 1312
 FT REPEAT 232 588
 FT REPEAT 830 1186
 FT METAL 395 395
 FT ACT_SITE 396 396
 FT METAL 399 399
 FT METAL 993 993
 FT ACT_SITE 994 994
 FT METAL 997 997
 FT CARBOHYD 59 59
 FT CARBOHYD 79 79
 FT CARBOHYD 116 116
 FT CARBOHYD 151 151
 FT CARBOHYD 165 165
 FT CARBOHYD 323 323
 FT CARBOHYD 514 514
 FT CARBOHYD 682 682
 FT CARBOHYD 700 700
 FT CARBOHYD 719 719
 FT CARBOHYD 765 765
 FT CARBOHYD 947 947
 FT CARBOHYD 1196 1196
 FT VARIANT 568 568
 FT SEQUENCE 1312 AA; 150947 MW; 9C13BB0529AD3755 CRC64;
 Query Match 31.1%; Score 1334; DB 1; Length 1312;
 Best Local Similarity 42.6%; Pred. No. 8.4e-85;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;
 QY 20 TIEQAKFLDKFNEAEDLFVQSSLASWYNTNTERNVONMNNAGDKWSAFLKEQSTL 79
 DB 649 TDEAKADFEVEYDRTAQVLLNEYAEANQWYNTNITIEGSKILLEKSTEVSNHTLKCTR 708
 QY 80 AQMYLQETQNLTVKQLQALQONGSSVLSSEKSKRLNTILNTWSTIYTGKVCNPDPQ 139
 DB 709 AKTFVSNFQNSRIKRLQNLDRAPLPPKELEYWQILLDMETIYSLSNICYTG-- 766
 QY 140 ECLLEPGLNETMANSLDYNERLWAWESWRSEVGKQLRPLYEYVVLKNEMARAHYEDY 199
 DB 767 TCMLEPDLTNMATSRYEELLNAWKSRRKVGRAILPPFKYVFNKTLAKNGYTD 826
 QY 200 GDYWRGDEVGVDYYSRGLIEDVHTTEELKPLVEHLYHAYVRAKLMNAYS-YISF 258
 DB 827 GDSWRSLYESNLE-----QDLKLYQELQPLVNLHAYVRRSLHRHYGSEYINL 876
 QY 259 IGCLPAHLGLDMWGRFTNLSLTVPGFGKPNIDVTDAMVDQAWDAQRIEAKFFVSF 318
 DB 877 DGPFAHLGLGNWAGTWSNIYDLVAPFSPAPNIDATEAMIKQGTTPRIFKEADNFTSL 936

QY 319 GLPNWTCQFWENSLMDPGNVQKAVCHPTAWDLKGG--DFRILMCTKYVTMDLFAHBMG 377
 DB 937 GLLPVPPFENKSMLEKPTDGRVYVCHPSAWDFYNGDKFRKQCTSVNMDLVIAHBMG 996
 QY 378 HIQYDMAYAAQPFLLRNGANGEGNEAVGEIMLSAATPKHLKSTGLLSPDFQEDNETIN 437
 DB 997 HIQYFMQYKDLPTVTFREGANPGFHEAIGDIALSVSTPKHLYSLNLSLSTE--GSGYEYDIN 1055
 QY 438 FLLKQALTYVGTLPPTYMLEKRWMMVFKGEIPKQDMKKMMKREIYVVPVPHDITY 497
 DB 1056 FLMKALDKIAFIPFSYILIDQWRVRVFDGSITTKENYQEWNSRLKYGGLCPPVPRSQGD 1115
 QY 498 CDPASLFHVSNDYSFIIRYVYTTLYQFOEALCOAAKHEGPLHKCDISNSTEAGQKLFNM 557
 DB 1116 FDPQKSFHPANVPYRVYFVFIQFOEALCAAAGTGPLHKCDIYQSKRAKGLADA 1175
 QY 558 LRLGKSEPTWLTALENVGAKNMVRPLNYEPLFTWLKQDK--NSFVGH--STDWSP 612
 DB 1176 MKLGYSKWPPEAMKLITQPNKNSASAMNYPKPLTEWLVLTENRRHGETLGHPEYNWAP 1233
 RESULT 5
 ACE_CHICK
 ID ACE_CHICK STANDARD; PRT; 1193 AA.
 AC Q10751;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Angiotensin-converting enzyme (EC 3.4.15.1) (Dipeptidyl
 DE carboxypeptidase I) (Kininase II) (Fragment).
 GN DCP1 OR ACE.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX TISSUE=Lung;
 RX MEDLINE=95110342; PubMed=7811282;
 RA Esther C.R., Thomas K.E., Bernstein K.E.;
 RT "Chicken lacks the testis specific isozyme of angiotensin converting
 RL Biochem. Biophys. Res. Commun. 205:1916-1921(1994).
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-l-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; L40175; AAA5554.1; -
 DR InterPro; IPR001548; Peptidase_M2.
 DR InterPro; IPR001130; Zn_MTpeptdse.
 DR Pfam; PF01401; Peptidase_M2; 2.
 DR ProDom; PD004184; Peptidase_M2; 2.
 DR PROSITE; PS00142; ZINC_PROTEASE; 2.
 KW Hydrolase; Metalloprotease; Carboxypeptidase; zinc; dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat.
 FT NON_TER 1 1
 FT DOMAIN <1 1156 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1157 1173 POTENTIAL.

```
FT DOMAIN 1174 1193 CYTOPLASMIC (POTENTIAL).
FT REPEAT 125 481
FT REPEAT 723 1079
FT METAL 288 288
FT METAL 289 289
FT ACT_SITE 292 292
FT METAL 886 886
FT METAL 887 887
FT ACT_SITE 887 887
FT METAL 890 890
FT CARBOHYD 11 11
FT CARBOHYD 60 60
FT CARBOHYD 216 216
FT CARBOHYD 407 407
FT CARBOHYD 447 447
FT CARBOHYD 485 485
FT CARBOHYD 513 513
FT CARBOHYD 555 555
FT CARBOHYD 575 575
FT CARBOHYD 658 658
FT CARBOHYD 1089 1089
SQ SEQUENCE 1193 AA; 137820 MW; 954472A18EA471C7 CRC64;

Query Match 30.6%; Score 1312; DB 1; Length 1193;
Best Local Similarity 40.4%; Pred. No. 2.5e-83;
Matches 264; Conservative 110; Mismatches 243; Indels 36; Gaps 10;

QY 22 EEOAKFLDKFHEDLFYQSSLASWNYNTNITENVNOMNAGDKWSAFLKEOSTLAQ 81
DB 544 EAOKEFLSEYNSTARVWNAATEASWEYNTNITDNKEVMEKLNAMSKHTIEYGMAR 603
QY 82 MYPLQBIQLTVKQLQALQNGSSVLSSEKSKRLNTINTMTYSTYTGKVCNPN---P 138
DB 604 QPDPSPQDETIVTRILNKLVSLEALPELDELKEYNTLLSDMETTYSVAKVCRENTEHP 663
QY 139 QECLLLEPGLEIMANSLDYNELNLAWSRSEVQKQLRPLVEEYVVLKNEARAHVED 198
DB 664 -----LDPDLTDILATSRDYNELLFAKGNWDASGAKIKDKYKRVYELSNKAALVNGYTD 718
QY 199 YGDYWRGDEVNGVDYDYSRGQIDVEDHTEETKPLHYELHAYVRAKLMNAY-PSYIS 257
DB 719 NGAYWSLSYETPTFE-----EDLERLYLQQLPVLNLHAYVRRALYKNYGAETHS 768
QY 258 PICCLPAHLGLDMGFRWNLVSLVPPFCQKNIDVTDAMVDQAMDQADRIKFAKEFFVS 317
DB 769 LKGPFAHLGLGNWAQSWNIEDLVMPFPDATKYVDATPAKQGGTTPKKNFEESDRFFTS 828
QY 318 VGLPNNMTQGFWSNMLTDPGNQKAYCHPTAWDL-GKGDFRILMCTKVTMDDFLTAHHEM 376
DB 829 LGLIPMPQEFWDKSMTEKPADGREVVCHASAWDFYNRKDFRKQCTVYVNDLLITVHHEM 888
QY 377 GHIOYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSLGSLSPDFQEDNETEI 436
DB 889 GHVOYFLQYMDQPISPRDGANGPGEHGAIGDMVALSVSTPKHLHSINLLD-QVTNEESDI 947
QY 437 NELLKQALTVIGTLPTTYMLEKRWNVFKGEIPKQDMKKKWKEMKEIEIVGVVEVPVPHDET 496
DB 948 NYLWSTALDKIAELPFGYLDQMRWKFVQGRKEDENYNOOBNRLUKYGLGCPVPRPSRD 1007
QY 497 YCDPASFLHVSNDYSFTRYRTLYQFOEALCQAAKHEGPHLKHCDISNSTEAGOKLFW 556
DB 1008 DFDGAKFHIPANVPYIRVFSVFIQFQHQALCKAAGHTGPHLTCDIYQSKRAGKLLGD 1067
QY 557 MLRLGKSEPTWLTALENVVGAKNNVRPLNLNYFEPLFTWL--KDONKNSFVGM-STDWSPY 613
DB 1068 AMKLGFSKPWPEAMQILGTGPNNMSAELMSYFEPLMTVLVKKNTENGEVLGWPEYSWTPE 1127
QY 614 ADOSIKVRISLKSALG-----DKAYENDNEMYLFRSSVAYAMQYFLKYK 659
DB 1128 AVTEFHAATDTADFLGMSVGTQKATAGW-----VLLALALVELITSIFLGVK 1175

RESULT 6
ACE_RAT
```

```
ID ACE_RAT STANDARD; PRT; 1313 AA.
AC P47820;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
DE (ACE) (Dipeptidyl carboxypeptidase I) (Kininase II).
DE DCP1 OR ACE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
RX MEDLINE=941121658; PubMed=8292044;
RA Koike G., Krieger J.E., Jacob H.J., Mukoyama M., Pratt R.E.,
RA Draai V.J.;
RA "Angiotensin converting enzyme and genetic hypertension: cloning of
RA rat cDNAs and characterization of the enzyme.";
RL Biochem. Biophys. Res. Commun. 198:380-386(1994).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=LEW/N; TISSUE=Lung;
RC Jafarian-Tehrani M., Listwak S., Barrientos R.M., Michaud A.,
RC Corvol P., Sternberg E.M.;
RA "Characterization of a missense mutation in the angiotensin
RA I-converting enzyme cDNA in exudative inflammation resistant F344/N
RA rats.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-l-xaa-xbb, when xaa is not pro, and xbb is neither
CC asp nor glu. Converts angiotensin I to angiotensin II.
CC -!- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME
CC IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY
CC FROM AN ALTERNATIVE START SITE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U03708; AAA82110.1; -
CC EMBL; U03734; AAA82111.1; -
CC EMBL; AF201332; AAG35597.1; -
CC InterPro: IPR001548; Peptidase_M2.
CC InterPro: IPR000130; Zn_MTpeptidse.
CC Pfam: PF01401; Peptidase_M2; 2.
CC PRINTS: PR00791; PEPDIPTASEA.
CC PRODom: PD004184; Peptidase_M2; 2.
CC PROSITE: PS00142; ZINC_PROTEASE; 2.
KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
FT SIGNAL 1 35 BY SIMILARITY.
FT CHAIN 36 1313 ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
FT ISOFORM.
FT DOMAIN 36 1265 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1266 1282 POTENTIAL.
FT DOMAIN 1283 1313 CYTOPLASMIC (POTENTIAL).
FT REPEAT 233 589
FT REPEAT 831 1187
FT METAL 396 396
FT ACT_SITE 397 397
FT METAL 400 400
FT ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT ZINC 1 (BY SIMILARITY).
FT ZINC 1 (CATALYTIC) (BY SIMILARITY).
```

FT	METAL	994	994	ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	995	995	2 (BY SIMILARITY).
FT	METAL	998	998	ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	44	44	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	60	60	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	80	80	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	117	117	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	152	152	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	166	166	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	324	324	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	515	515	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	583	583	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	701	701	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	720	720	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	766	766	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	948	948	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1197	1197	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARIANT	207	207	R -> K.
SQ	SEQUENCE	1313 AA;	150907 MW;	8CB5D0015F129591 CRC64;
Query Match 30.5%; Score 1310; DB 1; Length 1313;				
Best Local Similarity 42.0%; Pred. No. 3.9e-83;				
Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;				
QY	20	TIEBQAKFLDKFNHEADLFYQSLSASWNTNTNTEENVQNMNAGDKWSAFLKEOSTL	79	
Db	650	TDEAKANFEVEYDRTAKVLWNEAEANWHYNTNITIBGSKILLQKNKNEVSHPTLKYGTV	709	
QY	80	AQMYPLQEIQNLTVKLOLQALQONGSSVLSEKSKRLNTILNTMTSTYSTGKVCNPNQ	139	
Db	710	AKTFVSNFQNSIKRIKKVONVDRAVLPNPLEEYVQIILLDETTYSVANVCYTG--	767	
QY	140	ECLLEPGLNEIMANSIDNRLMAWESWRSEVGKQLRPLRYEEYVVLKNEAMRANHVEDY	199	
Db	768	TCLSEPLDTNIMATSRKYELLWVKRSWKVGRALPPFKVVDNSKIAKLNGVSDA	827	
QY	200	GDYWRGDEVNGVDGYSRQLIEDVEHTEETIKPLHYELHAYVRAKLNNAYS-YISP	258	
Db	828	GDSWRSSYESDDLE-----QDLKLYOELQPLYNLHAYVRSRHHYGEVINL	877	
QY	259	IGCLPAHLGDMGWFNTLYSLVTFGOKENIDVTAMVDQANDAOIRKPEAKFFVS	318	
Db	878	DGPFPALHLLGNWAGTNSIITDLVAPPSAPSIDATAMIKOGMTPRIFKEADNFFTS	937	
QY	319	GLPNMTQGFWSMLTDPGNVQKAVCHPTAMDGKG-DFRILMCTKVTMDDFLTAHHEMG	377	
Db	938	GLLPVPPEFWNKSMLKPTDGRVVCVCHASAMDFYNGKDFRIKQCTSVNMEELVIAHHEMG	997	
QY	378	HIQYDMAAQAQFLLRNCANCFHEANVEINSLSAATPKHLKSGILSLSPDFQEDNTEIN	437	
Db	998	HIQYFMQTKDLPVTRFGANCPGHEAIGDVALSVSTPKHLHSLNLSS-EGSYEHDI	1056	
QY	438	FLKQALITVGLTPTYMLEKRWMMFKGETPKDQWKKWEMKRETVGVVPEVPHDET	497	
Db	1057	FLMKALDKIAIFPSLIDQWRVRVDGSTITKENYQWWSRLKYQGLCPVPRSQGD	1116	
QY	498	CDPASLPHVSNDSYFIRYITRTLYQFQFQALCOAAKHEGLHKCDISNSTAGOKLFNM	557	
Db	1117	FDPGSKFHPANVPYIRYIFSIIFQFHEALCRAAGHTPLYKCDIYQSKAGKLLADA	1176	
QY	558	LRLCKSEPTWALENVGAKNMVRLNLYPEPLFTWLKQDNK--NSFVQW-STDWSP	612	
Db	1177	MKLGYSKQWPEAMKIITGQPNMSASAIMNYFKPLTEWLVTNRHGHGTGLGWPEYTWTP	1234	
RESULT 7				
ACE_RABIT	ACE_RABIT	STANDARD;	PRT;	1310 AA.
AC	P12822;	O02852;		
DT	01-Oct-1989	(Rel. 12, Created)		
DT	01-Nov-1997	(Rel. 35, Last sequence update)		
DT	30-May-2000	(Rel. 39, Last annotation update)		
DE	Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)			

DE	(ACE) (Dipeptidyl carboxypeptidase I) (Kininase II).
GN	DCP1 OR ACE.
OS	Oryctolagus cuniculus (Rabbit).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX	NCBI_TaxID=9986;
OX	[1]
RN	SEQUENCE FROM N.A.
RP	TISSUE=Lung;
RC	MEDLINE=92178960; PubMed=1311831;
RX	Thekkumkara T.J., Livingston W. III, Kumar R.S., Sen G.C.;
RA	*Use of alternative polyadenylation sites for tissue-specific
RT	transcription of two angiotensin-converting enzyme mRNAs.;
RT	Nucleic Acids Res. 20:683-687(1992).
RL	[2]
RN	REVISIONS.
RP	Sen G.C.;
RA	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL	[3]
RN	SEQUENCE OF 1-88 FROM N.A.
RP	TISSUE=Liver;
RC	MEDLINE=91139683; PubMed=1847388;
RX	Kumar R.S., Thekkumkara T.J., Sen G.C.;
RA	*The mRNAs encoding the two angiotensin-converting isozymes are
RT	transcribed from the same gene by a tissue-specific choice of
RT	alternative transcription initiation sites.;
RL	J. Biol. Chem. 266:3854-3862(1991).
RL	[4]
RN	SEQUENCE OF 34-55.
RP	TISSUE=Lung;
RC	MEDLINE=84051289; PubMed=6314908;
RX	Iwata K., Blacher R., Soffer R.L., Lai C.Y.;
RA	*Rabbit pulmonary angiotensin-converting enzyme: the NH2-terminal
RT	fragment with enzymatic activity and its formation from the native
RT	enzyme by NH4OH treatment.;
RL	Arch. Biochem. Biophys. 227:188-201(1983).
CC	-1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC	THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE
CC	VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC	-1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC	oligopeptide-I-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC	Asp nor Glu. Converts angiotensin I to angiotensin II.
CC	-1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME
CC	IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY
CC	FROM AN ALTERNATIVE START SITE.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X62551; CAA44428.1;
DR	EMBL; M58579; AAA31151.1; ALT_SEQ.
DR	PIR; A23455; A23455.
DR	PIR; S35484; S35484.
DR	InterPro; IPR001548; Peptidase_M2.
DR	InterPro; IPR000130; Zn_Mtpeptdse.
DR	Pfam; PF01401; Peptidase_M2; 2.
DR	PRINTS; PR00791; PEPDIPASEA.
DR	ProDom; PD004184; Peptidase_M2; 2.
DR	PROSITE; PS00142; ZINC_PROTEASE; 2.
KW	Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
KW	Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
FT	SIGNAL 1 33
FT	CHAIN 34 1310 ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
FT	ISOFORM.
FT	DOMAIN 34 1263 EXTRACELLULAR (POTENTIAL).


```
FT TRANSMEM 1264 1280
FT DOMAIN 1310
FT REPEAT 322 588
FT METAL 829 1185
FT ACT_SITE 395 395
FT METAL 396 396
FT METAL 399 399
FT ACT_SITE 993 993
FT METAL 996 996
FT CARBOHYD 59 59
FT CARBOHYD 79 79
FT CARBOHYD 151 151
FT CARBOHYD 323 323
FT CARBOHYD 449 449
FT CARBOHYD 513 513
FT CARBOHYD 681 681
FT CARBOHYD 699 699
FT CARBOHYD 718 718
FT CARBOHYD 946 946
FT CARBOHYD 1195 1195
FT CONFLICT 48 48
SQ SEQUENCE 1310 AA; 150405 MW; 04777FAB1798IDEA CRC64;

Query Match 29.9%; Score 1283.5; DB 1; Length 1310;
Best Local Similarity 36.8%; Pred. No. 2.7e-81;
Matches 275; Conservative 135; Mismatches 262; Indels 75; Gaps 17;

QY 7 LLLSLVAVTAAGSTI-----EQAKTFDKFNHEAEEDLFYQSSLASWYNTNI 54
DB 21 LLLLLLPPPAALTLDPLGLPGDFADEAGARLFASYNSSAEQVLFSTRSTAASWARDNI 80
QY 55 TEENVMNNAGDKWFAFKQSTLAQMYPLQEI-----QNTL---VKIQLAQALQNGSS 106
DB 81 TAENARQEE-----EALLSOFEAERAGKKAKELYDPWNTDTEFLRRIIGAVTLGPA 135
QY 107 VLSEDKSKRLNTLNTMTSTYGTGVNPDNPQECILLLEPGNEIMANSLDYNERLWAE 166
DB 136 NLPLAKROOYNLLSNISQIYSGKVCFFPKTASCWSLSDPLDNLNIIASSRYAMLLFAWE 195
QY 167 SWRSEVGKOLRPLYEVVVLKKNEMARAHVEDYGDYWRGDIYEVNGVDGYDSRGQLIEDV 226
DB 196 GWHNAVGIPLKPLYQETALSNEAYRQDGFSDTGATWRSWYDSPTFE-----EDL 245
QY 227 EHTFEIKPLYEHLHAYVRKALNNAY-PSYISPIGCLPAHLGLDMGWRWTNLYSLTVP 285
DB 246 ERIYHQLPEPLNLHAYVRRVLRHRRYGRYINLGRPIPAHLGLGNMAQSWESIYDMMVFP 305
QY 286 GQKPNIDVTAMDVDQAMDQRIKFAEKFFVSVGLPNNTQGTWENSLMDPNCNVOKAVCH 345
DB 306 PDKPNLDVTSYVQKGNATHMFRVAEEFFSLGLLPPMPPEFWAESMLEKPEDGREVWCH 365
QY 346 PTAWDL-GKGDPRILMCTVTWDDFLTAHHEGHIOYDMAYAAQPLLRNANGEGFHEAV 404
DB 366 ASAWDYNRDKFRKIOCTQVTMDLSTVHEHMGHVQYIYQYKQDPVSLRR-ANPGFHEAI 424
QY 405 GEIWSLSAATPKRLKISGLSPDFQEDNETEINFLKQALITVGLTFLPTMYLERKRWMYF 464
DB 425 GDVLLSVSTPAHLKIGLLD-HVTNDESDINYLKMALEKIAFLPPGYLVLDQNRGWVF 483
QY 465 KGEIPKQMKKWKEMKREIVGVVPHDEYCDPASLPHYSNDYSFTRYTRILYQXQ 524
DB 484 SGRTPSRVNFDDWYLRKYQICPPVVRNETHFDAGAKFHPISVTPYTRYFVSFVLQXQ 543
QY 525 FOEALCOAAHGEPLHKCDISNSTEAGOKLFNMLRGLKSEPTWTLALENVVGAKNMVRPL 584
DB 544 FHQALCWEAGHQPLHQCIDIYSTRAGAKLRAVLQAGCSRPPQWQEVULKDMVASDALDAQPL 603
QY 585 LNYFEPLFTWLKQDNKNS--FVGV-STDMSYADQSIKVRISLKSALG-----DKA 632
DB 604 LDYFQPVQWLOQNERNGEVLGWPEYQWRPPLPNPNYPEGIDLVTDABASRFSVEYDRS 663
QY 633 YE--WND-----NEMYLFRSSVAYAMROYFLKVKKNQMLFGEDYRVANLKRPRISNFF 684
```

```
DB 664 FOAVWNEYAEANNYNNTNITTEASKILLQKNQIAHNTLYG-----NWARRFDVSNF 716
QY 685 VTAP-----KNVSD-----IIPRTEVEK 702
DB 717 QNATSKRIIRKKVQDLQRAVLPVKEEE 743

RESULT 8
ACET_RABIT STANDARD; PRT; 737 AA.
ID ACET_RABIT
AC P22968;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Angiotensin-converting enzyme, testis-specific isoform precursor
DE (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (Kinase II).
GN DCP1 OR ACE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Testis;
RX MEDLINE=89380303; PubMed=2550457;
RA Kumar R.S., Kusari J., Roy S.N., Soffer R.L., Sen G.C.;
RT "Structure of testicular angiotensin-converting enzyme. A segmental
RT mosaic isoform."
RL J. Biol. Chem. 264:16754-16758(1989).
RN [2]
RP SEQUENCE OF 1-117 FROM N.A.
RX MEDLINE=91139683; PubMed=1847388;
RA Kumar R.S., Thekkumkara T.J., Sen G.C.;
RT "The mRNAs encoding the two angiotensin-converting isozymes are
RT transcribed from the same gene by a tissue-specific choice of
RT alternative transcription initiation sites."
RL J. Biol. Chem. 266:3854-3862(1991).
CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-l-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC -1- COFACTOR: BINDS 1 ZINC ION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME
CC IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY
CC FROM AN ALTERNATIVE START SITE.
CC -1- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.
CC -1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL
CC REGULATION BY ANDROGENS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; J05041; AAA31153.1; -
CC EMBL; M58580; AAA31152.1; -
CC PIR; A34402; A34402.
CC MEROPS; M02.004; -.
CC InterPro; IPR001548; Peptidase_M2.
CC InterPro; IPR000130; Zn_MTPeptide.
CC Pfam; PF01401; Peptidase_M2; 1.
CC PRINTS; PR00791; PEPTIDASEA.
CC ProDom; PD004184; Peptidase_M2; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
KW
```


KW Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing.
 FT SIGNAL 1 32
 FT CHAIN 33 737
 FT
 FT DOMAIN 33 690
 FT TRANSMEM 691 707
 FT DOMAIN 708 737
 FT METAL 419 439
 FT ACT_SITE 420 420
 FT METAL 423 423
 FT METAL 447 447
 FT CARBOHYD 108 108
 FT CARBOHYD 126 126
 FT CARBOHYD 145 145
 FT CARBOHYD 373 373
 FT CARBOHYD 622 622
 SQ SEQUENCE 737 AA; 83923 MW; FC43CC7655C3DCA CRC64;

 Query Match 29.9%; Score 1283; DB 1; Length 737;
 Best Local Similarity 40.8%; Pred. No. 1.3e-81;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

 QY 20 TLEEQAKTFLDKFHEAEDLYQSSLASWNTNITEE-----NVONNN--AGDKW 69
 DB 75 TDEARSFVEYDSFOAVNVEAEANNNTNITTEASKILLQKNQIANHTLYGNW 134
 QY 70 SAFLKEQSTLAQMYPLQIQNLTVKLOLQALQONGSSVLSBDSKRLNTILNTWSTYST 129
 DB 135 -----ARRDVSFNQATSKRIIKKVKQDQRAVLPKVEEYQILLDMETIYSV 184
 QY 130 GKVCNPDNPQECLELLEGLNEMANSNDYRLMAWESWRSEVQKRLPYEYVVLKNE 189
 DB 185 ANVCVRDQ--SCLQLEPLTLNMTATSRKYDELLWMTSWRDKVGRAILPYPKVVEFNK 242
 QY 190 MARAHYEDYGDYRGDEYVNGVDYDSRGLIEDVEHTFEERIKPLYEHLHAYVRAKLM 249
 DB 243 AARLNGYDAGDSWRSMTPTTLE-----QDLERLFOELQPLYNLHAYVGRALH 292
 QY 250 MAY-PSYISPIGCLPAHLGLDMWGRFTNLSLVPFGQKPNIDVTAMVDOANDQRIIF 308
 DB 293 RHYGAOHINLESPIPAHLLGNMAQTSNIYDLVAPFSPASTMDATAMTKQGTTPRMF 352
 QY 309 KEAEKFFVSVGLPNWGTGFWNSMLTDPGNVOKAVCHPTAWDLGK-DFRILMCTKYTMD 367
 DB 353 EADKFFLSGLPLVPPEFVFNKSMLEKPTDGRVNVCHASAWDFYNGDKFRKQCTVYME 412
 QY 368 DFLTAHMHGHIQDMAYAAOPFLLRNGANGEPHEAVEGTEMSLSAATPKHLKSLGLSPD 427
 DB 413 DLVVVHMHGHIQDMAYAAOPFLLRNGANGEPHEAVEGTEMSLSAATPKHLKSLGLSPD 472
 QY 428 FOEDNETEINFLKQALTIIVGTLPFTYMLKRWVFKGEIPKDKONMKWEMKREIVGV 487
 DB 473 -GGGYEHDINFLKMAKDIAFIPESYLVDEWRVRVFDGSGTITKENYQEWWSLRKQGL 531
 QY 488 VEPVPHDETCDPASLHVSNDSYFIRYTRTLVYFOFQALCAQAEGLPKHCKDTSNS 547
 DB 532 CPPAPRSGDPPGAKPHIPSSVPIRIFVFSFIQFHEALCAAGTGLTCDIYQS 591
 QY 548 TEAQKLFNMLRGLKSPWTLLENVYGNVAKNMYRPLINLYFELFTWLKDON--KNSFVG 605
 DB 592 KEACKRLADAMKLGYSKFWPEMKVITGQPNMSASAMNFKPLMDMLLTENGHRHGKLG 651
 QY 606 W-STDWSPYADQS 617
 DB 652 WPOYTWTPNSARS 664

 RESULT 9
 ID ACE_DROME STANDARD; PRT; 615 AA.
 AC Q10714; Q27572;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

30-MAY-2000 (Rel. 39, Last annotation update)
 Angiotensin-converting enzyme precursor (EC 3.4.15.1) (Dipeptidyl
 carboxypeptidase I) (Kininase II).
 ANCE OR RACE.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95293950; PubMed=7775412;
 RA Cornell M.J., Williams T.A., Lamango N.S., Coates D., Corvol P.,
 RA Soubrier F., Honeisel J., Lehrach H., Isaac R.E.;
 FT Cloning and expression of an evolutionary conserved single-domain
 FT angiotensin converting enzyme from Drosophila melanogaster.;
 RL J. Biol. Chem. 270:13613-13619(1995).
 RN [2]
 RP REVISIONS.
 RA Cornell M.J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CANTON-S;
 RX MEDLINE=96028519; PubMed=7547464;
 RA Tatei K., Cai H., Ip Y.T., Levine M.;
 FT Race: a Drosophila homologue of the angiotensin converting enzyme.;
 RL Mech. Dev. 51:157-168(1995).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE CONTRACTIONS OF THE HEART, GUT
 CC AND TESTES.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-[Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE AMNIOGROSOM DURING GERM BAND
 CC ELONGATION, SHORTENING AND HEART MORPHOGENESIS. EXPRESSED IN
 CC MIDGUT THROUGHOUT EMBRYOGENESIS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U25344; AAC02171.1; -;
 CC EMBL; U34599; AAC46902.1; -;
 CC MEROPS; M02.003; -;
 CC FlyBase; Fgn0012037; Ance.
 CC InterPro; IPR001548; Peptidase_M2.
 CC InterPro; IPR001330; Zn_MTpeptdse.
 CC Pfam; PF01401; Peptidase_M2; 1.
 CC PRINTS; PR00791; PEPTIDASEA.
 CC ProDom; PD004184; Peptidase_M2; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 FT Glycoprotein; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 615
 FT METAL 367 367
 FT ACT_SITE 368 368
 FT METAL 371 371
 FT CARBOHYD 53 53
 FT CARBOHYD 196 196
 FT CARBOHYD 311 311
 FT CONFLICT 48 51
 FT CONFLICT 141 141
 FT CONFLICT 293 293
 FT CONFLICT 486 486
 FT CONFLICT 533 533

QY 424 LSPDFQEDNETETINFLKQALTYVGLTPTTYMLEKRWVFKGEIPKQWKKWEMKRE 483
 DB 417 LK-NVSDNEARINQFLTALDKXIVFLPFAFTMDKYRWALFRQADKSEWCAFWKLE 475
 QY 484 IVGVVPEVPHDTCYDPAFLHVSNDYSFIRYRTLYQFQFQALCOAA-----KHG 537
 DB 476 YSGLEPPVVRTEKDPADAKYHVSADVEYLYLVFIIFQFQYKSACTAGCYVPMOTEY 535
 QY 538 PLHKDCISNSTEAGOKLFNMLRLGKSEPTWTLAENNVCAKNNVRLNLYFEPLFTWLK 596
 DB 536 PLDNCDIYSGKAGKLFENMLSLGASKPWPDALEAFNGERTWTKATAYEPLRVWLE 594

RESULT 11

YFWA_BACSU
 ID YFWA_BACSU STANDARD; PRT; 501 AA.
 AC P50848;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 58.2 kDa protein in KDGT-XPT intergenic region.
 GN YFWA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCHI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / MARBURG;
 RX MEDLINE=96349105; PubMed=8760912;
 RA SOROKIN A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
 RA Serror P.;
 RT "Sequence analysis of the Bacillus subtilis chromosome region between
 RT the sera and kdg loci cloned in a yeast artificial chromosome.";
 RL Microbiology 142:2005-2016(1996).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M32 (ZINC
 CC METALLOPROTEASE).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L47838; AAB38482.1; -;
 DR EMBL; L77246; AAA96610.1; -;
 DR EMBL; Z99115; CAB14125.1; -;
 DR MEROPS; M32.001; -;
 DR Subtilist; BG11458; YFWA.
 DR InterPro; IPR001333; Peptidase_M32.
 DR InterPro; IPR000130; Zn_MTPeptidse.
 DR Pfam; PF02074; Peptidase_M32; 1.
 DR PRINTS; PR00998; CRBOXYPTASET.
 DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
 KW Hypothetical protein; Hydrolase; Metalloprotease; Zinc;
 KW Complete proteome.
 FT METAL 265 265 ZINC (CATALYTIC) (POTENTIAL).
 FT ACT_SITE 266 266 POTENTIAL.
 FT METAL 269 269 ZINC (CATALYTIC) (POTENTIAL).
 SQ SEQUENCE 501 AA; 58174 MW; A7489BBAFEA38F82 CRC64;
 Query Match 3.1%; Score 135; DB 1; Length 501;
 Best Local Similarity 20.9%; Pred. No. 0.056;
 Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;
 QY 180 YEEYVWL-KNEMA--RANHVEDYGG-----DYWRGDYEVNGVDGYDS 218
 DB 103 YKEYVILCSKAETAAWEAKGSDFLSPYLEQLIEPKRFTITW-----GVQ-- 150
 QY 219 RGQLIEDVEHTFEIKPLYEHLHAYVAKLNNAYPSVISPGCLPAHLGLDMGRFTWNL 278

DB 151 -----EHPYDALLDLFE---PGVTVKYLD-----QLPAEL 177
 QY 279 YSLTVFP-----GQPNIDVTDMVQDAQRIKFAEKFFVSVGLPNMTQGFWENS 331
 DB 178 KEALIPLVKQVTASGNKP-----DTSFITKAPFKKOKELSLYFLQELGYDF----- 224
 QY 332 MLTDPGNVOACVCHPTAWDLGKGFRLMCTKVTMDDFLTA---HEEMGHIOYD--MAY 385
 DB 225 ---DGGRLDEV-HPFAATTLNRGDRV--TTRYDERKDFRTAIFGTHECCHAIYEQWIDE 278
 QY 386 AAQPFLLRNGANEGFHEA-----VGE-----IMLSAATPKHLKSIIGLLSPDF- 428
 DB 279 ALSGTNLSDGASMGIHESQSILFVENFGRNKHFWTPYKKIQEASPVQFKDISL--DDFV 336
 QY 429 QEDNETEINFLKQAA--LTIVGLTPTTYMLEKRWVFKGEIPKQWKKWEMKREIVG 486
 DB 337 RAINEKSPSIRVEADELTYPHIIIRYEIEK---AIFSNEVSVEDLPSLWKNQYQDYL 393
 QY 487 VPEVPHEDETCYDPAFLHVSNDYS-FIRYRTLYQFQFQALCO-----RAKHGGL 539
 DB 394 I---TPQTDAGILQDVHWAGGDFGFPFVALGYMTAQLKQMLDLDFEDALLERGEF 450
 QY 540 HKDCISNSTEAGOKLFNMLRL-KSEPWTALLENVVGAKNNVRLNLY 587
 DB 451 HPIK-----QWLTEKVHIGKRRKPLDIKDATG-EELNVRYLIDY 490
 RESULT 12
 GLGB_BACSU
 ID GLGB_BACSU STANDARD; PRT; 627 AA.
 AC P391118;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
 DE enzyme).
 GN GLGB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCHI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=94195107; PubMed=8145641;
 RA Kiel J.A.K.W., Bosls J.M., Beldman G., Venema G.;
 RT "Glycogen in Bacillus subtilis: molecular characterization of an
 RT operon encoding enzymes involved in glycogen biosynthesis and
 RT degradation.";
 RL Mol. Microbiol. 11:203-218(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98048467; PubMed=9387221;
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
 RT "Sequencing and functional annotation of the Bacillus subtilis genes
 RT in the 200 kb rmb-dnaB region.";
 RL Microbiology 143:3431-3441(1997).
 CC -!- FUNCTION: CATALYSES THE FORMATION OF THE ALPHA-1,6-GLUCOSIDIC
 CC LINKAGES IN GLYCOGEN BY SCLSSION OF A 1,4-ALPHA-LINKED
 CC OLIGOSACCHARIDE FROM GROWING ALPHA-1,4-GLUCAN CHAINS AND THE
 CC SUBSEQUENT TRANSFER OF THE OLIGOSACCHARIDE TO OTHER PARTS OF
 CC ALPHA-1,4-GLUCAN CHAINS.
 CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
 CC glycogen.
 CC -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
 CC -!- INDUCTION: EXPRESSED EXCLUSIVELY ON MEDIA CONTAINING CARBON
 CC SOURCES THAT ALLOW EFFICIENT SPOULATION.
 CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z25795; CAA81040.1; -
DR EMBL; AF008220; AAC00214.1; -
DR EMBL; Z99119; CAB15076.1; -
DR PIR; S36624; S36624.
DR Subtilist; BG10907; g1g8.
DR InterPro; IPR000461; Alpha_amlase.
DR InterPro; IPR004193; isomylase.N.
DR Pfam; PF00128; alpha-amlase; 1.
DR Pfam; PF02922; isomylase.N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT ACT_SITE 309 BY SIMILARITY.
FT ACT_SITE 352 BY SIMILARITY.
FT ACT_SITE 420 BY SIMILARITY.
SQ SEQUENCE 627 AA; 73665 MW; 6480A553B6767BA CRC64;

Query Match 2.9%; Score 125; DB 1; Length 627;
Best Local Similarity 18.6%; Pred. No. 0.38; 225; Indels 264; Gaps 37;
Matches 131; Conservative 86; Mismatches 86

QY 49 NYNTNITENVMNAGDKWSAFL-----KEQSTLAQYPLQETQNTLVKQLQALQON 103
DB 56 DNSNSGEHVHRYNDNCIWTLPFGIGEKER-----YKYEIVTN-----N 97
QY 104 GSSVLSEKSKRLNTILMTSTYTGKVCNPDNPQECILLLEPGLNETMAN--SLDYNER 161
DB 98 GRIKLKADP-----YAIYS-----EYRNPATSLTDLE 125
QY 162 LAAWSEWSEVSKQLRPLEYEVVVLKNEARAHVEDYDGYVNGVDYDSRGQ 221
DB 126 GYSWDQKQKKQAKTYEKVPFI-----YELHLSWK-----KHSDGRHSYKE 171
QY 222 LIEDV-----EHTFE--EIKPLYEHLHAYVRAKLMNAYPSYISPGCLPAHLGLDMWGRF 274
DB 172 LSQTLIPYIKKHGTHIELLPVE--HPYDRS-----WQY 205
QY 275 WTNLYSLVPPQCKNIDVTDAMVQA-----WDAQRIKAEKFFSVSLPNT 324
DB 206 GTGYSPSTRFG--PPHDLM--KFVDECHQNTGVILDWPGHFKDAGLNFDEGP--- 259
QY 325 QGFWENSLTDPGNYQKAVCHPTAMDGLKGDPRILMCKTKVMTDDFLTAH----HEMGIHQ 380
DB 260 --LYEYKEERDEN-----NLWGTANFDL--GKPEVHSFLSNALYWAEFYH- 303
QY 381 YDMAYAAQPFLLRNGANEGFH--EAYGEIMSLSAATPKHLKSTGLLSLPQEDNETEINFL 439
DB 304 -----DGRFVDAVANILYWPNDQDERH-----TNPYAVDFL 333
QY 440 LQAQTLVGTLPFTYML--EKRWVVFKEIPK-----QWKKWEMKREIVGVVEPV 491
DB 334 KKLNTQMRREAYPHVMIAEDSTEWFOVYGAEEGLGPHYKWMGW---MNDVLYMETP 390
QY 492 PHDETCDPASLHVSNDSYFIRYTRTLTYQFOQALCQAAKHGCLPKCKDISNSTRAG 551
DB 391 PEERRHC-----HQLISFSL-----YAFSEHFVLPFSHDEVV-----G 425
QY 552 OKLFNMLRLCKSEPWTALENVVGAKNMVRPLNY-----PEPLFTWLK 596
DB 426 KK--SLLNKMPGDYW-----OKFAQYRLLGYMTVHPGKKLIFMGSEFAQFDW-K 473
QY 597 DQKNKSFVGNSTWSPYADQSI-----KYRISLKSALGDKAVENW---NEMVL 642
DB 474 DTEQ---LDFWLSDFPMHQKASVFTQDLRLFFYQSKILYEHDRHQSEIDVHDEQSI 530
QY 643 FRSSVAYAMROYFLKVKNOIMILFGEHDVAVANLKPRISENFVFTAP 688
DB 531 F-SFIRYGQKH-----GEALVICNFTPVVYHOYDVGVP 563

RESULT 13
EPIB_STAEP STANDARD; PRT; 986 AA.
AC P30195;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Epidermin biosynthesis protein epib.
GN EPIB.
OS Staphylococcus epidermidis.
OG Plasmid ptu 32.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RU 3298 / DSM 3095;
RX MEDLINE=92155237; PubMed=1740156;
RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
Goetz F., Entlian K.-D.;
RT *Analysis of genes involved in the biosynthesis of lantibiotic
epidermin*;
RL Eur. J. Biochem. 204:57-68(1992).
CC -!- FUNCTION: INVOLVED IN THE POSTTRANSLATIONAL MODIFICATION OF THE
LANTIBIOTIC EPIDERMIN.
CC -!- SUBCELLULAR LOCATION: POSSIBLY ASSOCIATED WITH, AND ANCHORED TO,
THE CYTOPLASMIC SIDE OF THE MEMBRANE.
CC -!- SIMILARITY: TO B.SUBTILIS SPAB AND L.LACTIS NISB.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X62386; CAA44253.1; -
DR Plasmid.
KW Plasmid.
SQ SEQUENCE 986 AA; 116708 MW; 343CE98526847850 CRC64;
Query Match 2.9%; Score 123; DB 1; Length 986;
Best Local Similarity 18.2%; Pred. No. 1;
Matches 153; Conservative 120; Mismatches 258; Indels 310; Gaps 40;
QY 1 MSSSSWLLLSILVATAAQSTIEQAKTFLDKFHE-----AEDL--FYQSSLASWNY 50
DB 320 ISEAYILNLLSPNHFGTKIRNYHEFFMDKYGFQQLVNLKQLLSDINGFGYPKKDSYF 379
QY 51 NTNITEENVOMNAGDKWSAFLKEQSTLAQYPLQEQNLTKVLQALQOQSSVLS 110
DB 380 SNNI-----AFLKEK-----YLLATONNSHIE 404
QY 111 DKSKRLNTILMTSTYTGKVCNPDNPQECILLLEPGLNETMANSLDYNERLAWESWRS 170
DB 405 NDVANLEK--NTVSKINA-----PVSTEI-----YSHIF----- 433
QY 171 EVGKQLRPLYEYVVLKNEARAHNEYEDYDGYVNGVDYDSRGOLIEDVHTF 230
DB 434 --GNSIKG--YEDFAVISPLGSPNAGATGFR--TGNEFIK-----KKQLQKEIVH-- 480
QY 231 EEIKPLYEHLHAYVRAKLMNAYPSYISPGCLPAHLGLDMKGREWTNLYSLTVPFGOKPN 290
DB 481 -----HYANYMNGLEISOLNEGLSRNNTLLNN--NRIYNTCLNLNP---KSD 527
QY 291 IDVTDAWQDAQRIKFE---AEKFFSVVGLPNTQG-----FWENSMLTDPG 337
DB 528 IDINDIFGATFNKLYLYSEKHSRIYFVSNMSEYFEGSLEYKFLREISPEKTFIOPI 587
QY 338 N-----VQKAVCHPTAMDGLKGDPRILMCKTKVMTDDFLTAHHEMGIQYDM 383
DB 588 TEEGIDSLPFCPRIIYKNIILKPAWKINSEMFSE---TENWLNRFATI--REKWHIPKDV 643

```
QY 384 AYA-----AQPELLRNGANEGFHEAVGSGIMS----- 409
DB 644 IIAFGDNRLLNLLNDKHLILKELKHGRIRILESPINESNNRMLDIVPLPKKISL 703
QY 410 --LSAATPK-----HLK---SIGLLSPDFQEDNTEINFLLKQALTIVGTLPFTYMLE 457
DB 704 KEQSFIPKRNKRNKHNKADWFSIHLSIPKTYQD-----NFTQDVLPLPITELKVNPNFN 758
QY 458 KRWEMVFKGEIPKQDQMKKWMKMRIGVGVPEPHDETYCDPASLFHVSNDYSFIYY- 516
DB 759 KFYIYFKED--EDPIK--LRLRE-----DEDY-----SOIYFIKNNK 794
QY 517 -----TRTYQFQFBALCOAAKHEGPLHKCDISNSTEAGQKLFNMLRGKSEPTLALE 571
DB 795 DYCLLNSELYSDIVYVPEVRYGPHVIEDIENFMYDSLL--SINIQSE-FKIPKE 851
QY 572 NVYGANKMVRPLLYNFEPLFTWLKDONKNSFVGSWSDWSPYADOSIKVRISLKSGALGDK 631
DB 852 FIVA-----ISIDFLDLYE-----INKSEKEILINNA--ED 882
QY 632 AYEWNDNEMYLFRSSVAYAMROYFLKXNQMLFGEEDVRVANLKPRIISNFVTPAKNV 691
DB 883 LYRSND-----IREY-----KNLLAKL-----TNPKNND 905
QY 692 SDIIP-----RTEVEKAIRMSRSRINDAF-----RLNDNSLEFLGIQ 728
DB 906 YEILKKEFPNLHFLFNKISILENLKKTQKSLYTSRSRIISFIHMRGN-----RIFGIN 961
QY 729 P 729
DB 962 P 962

RESULT 14
BGAL_BACME STANDARD; PRT; 1034 AA.
AC 052847;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Beta-galactosidase (EC 3.2.1.23) (lactase).
GN BGAL.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 319;
RA Strey J.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC galactose residues in beta-D-galactosides.
CC -|- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ000733; CAA04267.1;
CC HSP: P00722; I8GL.
CC InterPro: IPR004200; Bgal_small_C.
CC InterPro: IPR004199; Bgal_small_N.
CC InterPro: IPR001649; Glyco_hydro_2.
CC Pfam: PF02930; Bgal_small_C; 1.
CC Pfam: PF02929; Bgal_small_N; 1.
CC Pfam: PF00703; Glyco_hydro_2; 1.
CC Pfam: PF02836; Glyco_hydro_2_C; 1.
```

```
DR Pfam: PF02837; Glyco_hydro_2_N; 1.
DR PRINTS: PR00132; GLHYDRASE2.
DR PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 481 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 547 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 1034 AA; 118673 MW; 38644C9A649A15E9 CRC64;

Query Match 2.88; Score 121; DB 1; Length 1034;
Best Local Similarity 19.28; Pred. No. 1.5;
Matches 104; Conservative 66; Mismatches 165; Indels 206; Gaps 30;

QY 29 LDKFNHEADLFYSSSL-----ASWNYNTNITEENVQNMN---NAGDKKSAFLKE 75
DB 260 LNYFEDHQDTTFEVMLYDANAQEVLPQATNLSVDSQRTVSLRTHIKSPAKWSA----- 315
QY 76 QSTLAQMYPLQEIOML-TVKLQALQOQSSVLSEDSK---RLNTILNTMTSYSTGK 131
DB 316 -----ESPPLYTLVL--SLKNAAGSIITETESCKVGFRTEIKNGLMI--NGK 359
QY 132 -----VCNPDNPOECILLLEPGLMEIMANSLDYNERLW----- 163
DB 360 RIVLRGVNRHEFDSYKGRAGITREDMIHDLMLKOHINAVRTSHYPNDSVWYELCNEYG 419
QY 164 -----AW-----ESWRSEVGKQLRPLYEE----- 182
DB 420 LYVIDETNLEHTGWTYLYQEGEKAVPGSKPENKENVLDRCRMYERDKNHPISIIISLG 479
QY 183 -----YVVLK-NEMARANHYEDGYWRGDEVGVNGVDGYDYSRGQLIEDVEHT 229
DB 480 NESTGGENFQHMVTFEKKEDSTRLVHYE--GIFHHRDYDASDIESTMYVKP---ADVE-- 532
QY 230 FEEIKPLYEHLHAVRAKLMAVSYISPIGCLPAHLGLDMGRFWMVNLVSLVYPPGQKP 289
DB 533 -----RYALMNPKKPYIL--CEYSHAMGNSCG---NLYKYWELFDQYP 570
QY 290 NID---VTDMVDAQDAORIFKEAE--KFFVSVG-----LPNMTQGFWEWSMLTDGCV 339
DB 571 ILQGGFIWD-----WKDQALQATADGTSTLAYGGFGDTPN-DGNFCGNGLLIFADGTA 623
QY 340 QKAV-----CH-PTAW---DLGKGFRLMCTKVTMDDFLTAHHEMGHI-----QYDMAY 385
DB 624 SPKIAEVKKCYQPVKWTAVDPAKGFV-----QNKHLFTNLNAYDFVW 667
QY 386 AAQPFLLRNGANEGFHEAVGSGIMSLSRAATPKHLSIGLLSPDFQEDNTEINFLLKQALT 445
DB 668 TVE-----KNG-----ELVEKHASLLNVPADGTDELTLSTPLYEQENETD-EFVLTLSLR 716
QY 446 I 446
DB 717 L 717

RESULT 15
YAMB_SCHPO STANDARD; PRT; 3655 AA.
AC Q10064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 420.8 kDa protein Clf5.11C in chromosome I.
GN SPAC1F5.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S.; Odell C.; Churcher C.M.; Barrell B.G.; Rajandream M.A.;
RA Walsh S.V.;
```

```
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: STRONG, TO YEAST YHR099W.
CC -!- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68136; CRA92239.1; -
DR InterPro; IPR003151; FATC.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR004003; PI3_PI4_kinase.
DR InterPro; IPR001440; TPR.
DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR Pfam; PF00454; PI3_PI4_kinase; 1.
DR SMART; SM00146; PI3Kc; 1.
DR PROSITE; PS02090; PI3_4_KINASE_3; 1.
KW Hypothetical protein; Transferase; Kinase.
FT DOMAIN 3324 3655 PI3K/PI4K (BY SIMILARITY).
SQ SEQUENCE 3655 AA; 420774 MW; 50475E3FC2124A CRC64;

Query Match          2.8%; Score 119.5; DB 1; Length 3655;
Best Local Similarity 17.0%; Pred. No. 11;
Matches 93; Conservative 93; Mismatches 197; Indels 165; Gaps 23;

QY 4 SSWLLLSLVAVTAAGSTIEEQAQT---FLDKPNHEADLFY-----QSSIASNNYNTNITE 56
DB 2633 SDW-----STEQDTLEKATKSLSPSTSLRRHTADALLYLNKTQRKMGSVTEFSRIID 2684

QY 57 ENVQNNNAGDKWSAFLE--QSTLAQMPLOEIONLVKLQALQOQSGSVLSEDKSK 114
DB 2685 ECMQFSLR---RWQQLPKRVYQSHVSLHHFQEIHEL-----QFAFGIYSQLNDT 2731

QY 115 RLNTILNTMTSYSGKVCNPDPNQCEILLPEGLNEIMANSIDYNERLWA-WESWRSEVG 173
DB 2732 NIHNIDNKL-----RDIKVLQGWRELRPNWWD-DIDWSDLIARQSVF 2775

QY 174 KQLRPDYEEYVVLKEMARANHEDYGDYWRGDEYVNGVDYDYSRGQLIEDVEHTFEEI 233
DB 2776 KSINKVFLPLVSIQAQSTNKSNTNSVLYRGYHEL----- 2811

QY 234 KPLYEHLHAYVRAKLMNAVPSYISPGCLPAHLGLDMGRCFWTNLYSLTVPPGQKPNIDV 293
DB 2812 -----AWIINRF-AHVARVHHLPEVCINQL-----TKIYTL-----PNIEI 2846

QY 294 TDAMVDQAMDQRIKFAEKFFVSVGLPNMT-----QGFWENSMLTDPG 337
DB 2847 QEAFKLRLQEAECHEVSEPMQGLGVINNTNLMYFRNOKAEFFTLKGMFON-RLGEKD 2905

QY 338 NVQKAVCHPTANDLKGDPRIILMCKYVMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGAN 397
DB 2906 EANAQAFATAVQIDIGSG-----KAWSEWG-LYHDELFOANPQEIHHACN 2948

QY 398 --EGFHEAVGEIMSLSAATPKHLKSLGLSPDFQEDNETEINFLKQALTIVGTLPFTYM 455
DB 2949 AVSCFLQA--SSLSNSSNAPLLTRVLNLLSVDDSDSGSVSEV-----VSSPKSE 2995

QY 456 LEKRWMMVFKEITPDQMMKWKWKREIVGVVPEPHDPTCYCDPASLPHVSNDY-SFIR 514
DB 2996 IPTNNWIPF---IPQ-----LLSALSHRSHSIHARAILIQIAKTYPQSLH 3036

QY 515 YVTRTLXQ 522
DB 3037 FQLRTAYE 3044
```